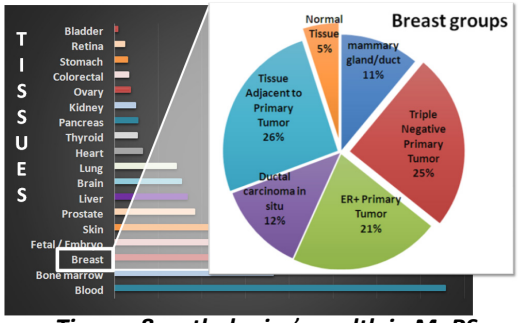
The MaRS (Matrix of RNA-Seq) Project

Fabien Pierrat[1], **L. Manchon**[1], **R. Bruno**[1] , **D. Piquemal**[1]
B. Cirou[2], **V. Cameo Ponz**[2], **F. Daumas**[2]

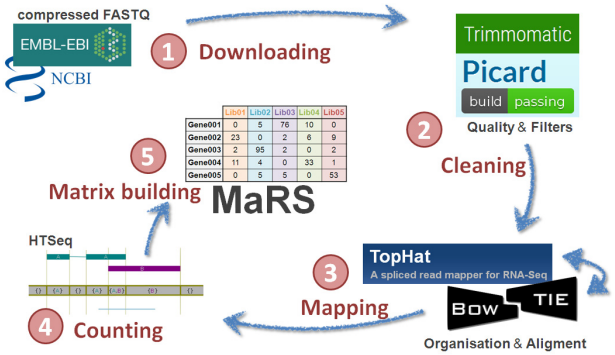
[1] *ACOBIOM, Montpellier, France*
[2] *CINES, Montpellier, France*

Recent years have seen a dramatic **increase** in the **amount** of genomic and **transcriptomic data** produced by laboratories around the world. The aim of the **MaRS research program** is to **collect** and to allow the **comparison** of these data. MaRS is focused on the **RNA-Seq** method, which reflects the expression of the genes in a specific condition.



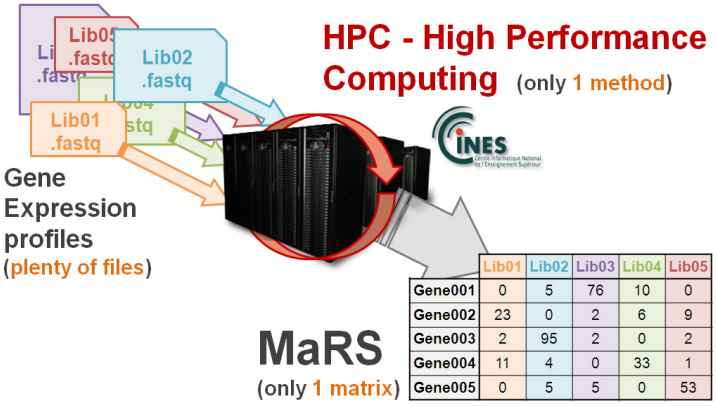
Tissues & pathologies' wealth in MaRS

27000 Human RNA-Seq profiles are selected for the project. In collaboration with the CINES, all these data are **collected** and **computed** with a single **method** to allow their **compilation** in **one matrix**: MaRS.
It represents a **huge amount of data** and requires the use of High performance computing (**HPC**) on the cluster ***Occigen*** (CINES): 120 To of compressed data downloaded and 1.2 million hours/core consumed.



High Performance Computing toolchain

RNA-Seq method is used in a wide variety of applications like identifying **disease-related** genes, analysing the effects of drugs on **tissues** or providing insight into disease pathways. The RNA-Seq is widely used to identify **gene expression patterns** associated with tumor formation.
MaRS represents a fantastic tool for the discovery and the validation of **Biomarkers**.



The MaRS project

Many **challenges** were encountered and solved during the project:
- the **download** of the very large amount of data
- the **installation** and the **settings** of the software's in the cluster
- the **optimization** of the compiler and the packages
- the **constraints** in jobs duration and jobs limitations.

➔Next? The Exploration of MaRS! Implementation of an advanced search tool to query the matrix in order to highlight biomarkers.
➔Extend work to other species: Mouse.

Corresponding author:
Fabien PIERRAT
pierrat@acobiom.com

In partnership with:

ACOBIOM

1682, rue de la Valsière / CS 77394 – Cap Delta
Biopôle Euromédecine II / 34184 MONTPELLIER Cedex 04
Phone: +33 (0)467 419 748 / Fax: +33 (0)467 457 726
info@acobiom.com / www.acobiom.com

CINES Centre Informatique National de l'Enseignement Supérieur